AD-A140 684 APPROXIMATING THE LEVEL PROBABILITIES IN ORDER RESTRICTED INFERENCE: THE..(U) MISSOURI UNIV-ROLLA DEPT OF MATHEMATICS AND STATISTICS F T WRIGHT ET AL. APR 84 1/1

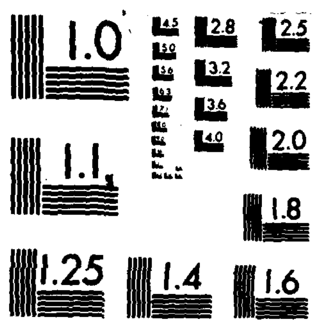
UNCLASSIFIED N00014-80-C-0322 F/G 12/1 NL

END

DATE FILMED

6-84

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

AD-A140 684

APPROXIMATING THE LEVEL PROBABILITIES IN ORDER RESTRICTED INFERENCE: THE SIMPLE TREE ORDERING[1]

by

F. T. Wright and Tuan Tran
Department of Mathematics and Statistics
University of Missouri-Rolla
Rolla, MO 65401

Some Key Words: Order restricted inference, comparisons with a control, simple tree orderings, level probabilities, chi-bar-squared distributions

(1) This research was sponsored by the Office of Naval Research under contract N00014-80-C0322.

SUMMARY

Distribution theory for the likelihood ratio statistics for the comparison of several treatments with a control is discussed. These test statistics account for prior information that the treatments are at least as effective as the control. Their null distributions are mixtures of chi-squared or beta distributions and the mixing coefficients, which are "level probabilities", are intractable for even a moderate number of treatments and unequal weights, which are typically the sample sizes. The distribution corresponding to equal weights is considered as an approximation and an approximation based on the pattern of large and small weights is developed. Both are adequate for moderate variations in the weights, but the approximation based on patterns in the weight set is considerably more accurate for larger variations.

1. INTRODUCTION

We consider the situation in which one wishes to compare several treatments with a control or standard when it is believed that the treatments are at least as effective as the control. For example, in a drug study, several drugs may be compared to a zero dose control.

Assuming that the observations are normally distributed, let H_1 denote the hypothesis that $\mu_1 \leq \mu_i$ for $i = 2,\ldots,k$, where μ_1 is the control mean and the μ_i for $2 \leq i \leq k$, are the treatment means. Barlow, Bartholomew, Bremner and Brunk (1972) discuss testing homogeneity, ie. H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$ versus an arbitrary partial ordering on the μ_i and call the ordering specified by H_1 a simple tree ordering. Robertson and Wegman (1978) consider tests with the partial order restriction the null hypothesis. In both of these situations, the distribution of the likelihood ratio test (LRT) under H_0 is a mixture of chi-squared or beta distributions depending on whether the variances are known or not. The mixing coefficients are the probabilities, under H_0, that the maximum likelihood estimates (MLEs) of the μ_i, subject to the partial order restriction, have a specified number of distinct values. These coefficients, which are also called level probabilities, depend upon the precisions of the sample means. The precisions will be referred to as weights. We will see that, even in the case of a simple tree ordering, the level probabilities can be very tedious to compute for unequal weights and moderate k.

Dunnett (1955, 1964) proposed a test of H_0 versus H_1-H_0, that is H_1 but not H_0. For the case of unknown variances, the test statistic is the maximum of appropriate pairwise t-test statistics. Robertson and Wright (1983b) studied the power functions for this test, the LRT and the one-sided t test which compares the control mean with the pooled treatment means. While Dunnett's test and the t test are more powerful in some subregions of the alternative space, the LRT maintains a more reasonable power throughout H_1.

We consider approximations for the level probabilities needed to implement the LRT for a simple tree ordering and unequal weights. For a total order such as $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_k$, Siskind (1976) and Grove (1980) observed that the level probabilities are robust to moderate variations in the weights and conjectured that, except in extreme cases, the level probabilities associated with equal weights should provide reasonable approximations. Robertson and Wright (1983a) studied their conjecture in the totally ordered case and found that for many practical applications this approximation is satisfactory. They also found that the quality of the approximation depends on the pattern of the larger and smaller weights, and they developed an approximation based upon such patterns. Following their work we study the equal-weights level probabilities as an approximation and develop an approximation based on the pattern of large and small weights for the simple tree ordering.

In Section 2, the level probabilities are studied as functions of the weights. Upper and lower bounds for the tail probabilities of the LRT statistics are determined and these bounds are shown to be sharp. In Section 3, the accuracy of the equal-weights approximation is considered for particular types of weight sets, as well as randomly generated weight sets. As in the totally ordered case, it is found that for many practical situations the equal-weights approximation is adequate. However, for large variation in the weights and certain patterns of large and small weights, there may be considerable error in this approximation. In Section 4, the limiting values of these mixing coefficients are obtained for each possible weight set in which only two values are assumed with one value constant and the other approaching infinity. Using these results, an approximation is proposed which outperforms the equal-weights approximation for weight sets with a large amount of variation.

2. THE LEVEL PROBABILITIES

Suppose X_{ij}, $j = 1,\ldots,n_i$ and $i = 1,\ldots,k$ are independent random samples from k normal populations with means μ_i and variances $\sigma_i^2 = a_i^2 b$ with a_i^2 known and $b = 1$ if known. Let $\bar{X}_1,\ldots,\bar{X}_k$ denote the sample means, let $w = (w_1,\ldots,w_k)$ with $w_i = n_i/a_i^2$ denote the vector of weights and let $P(\ell,k;w)$ denote the probability, under H_0, that $\bar{\mu} = (\bar{\mu}_1,\ldots,\bar{\mu}_k)$, the MLE of $\mu = (\mu_1,\ldots,\mu_k)$ subject to the restriction H_1, has exactly ℓ distinct values.

For the case of known variances, the LRT statistic for H_0 versus $H_1 - H_0$ is $T_{01} = \sum_{i=1}^k w_i(\bar{\mu}_i-\hat{\mu})^2$ where $\hat{\mu} = \sum_{i=1}^k w_i\bar{X}_i/\sum_{i=1}^k w_i$ and under H_0,

$$\bar{\chi}_w^2(t) \equiv pr(T_{01} \geq t) = \sum_{\ell=2}^k P(\ell,k;w)pr(\chi_{\ell-1}^2 \geq t) \text{ for } t > 0. \quad (1)$$

The LRT statistic for H_1 versus H_2: $\mu_i < \mu_1$ for some i is $T_{12} = \sum_{i=1}^k w_i(\bar{\mu}_i-\bar{X}_i)^2$. Within H_1, H_0 is least favorable and under H_0,

$$pr(T_{12}\geq t) = \sum_{\ell=1}^{k-1} P(\ell,k;w)pr(\chi_{k-\ell}^2\geq t) \text{ for } t>0. \quad (2)$$

The $\bar{\chi}^2$ distributions in (1) and (2) arise as approximations when considering multinomial parameters (cf. Robertson (1978)), one-parameter exponential families (Robertson and Wegman (1978)), Poisson intensities (Magel and Wright (1984)) and nonparametric tests (Shirley (1977) and Robertson and Wright (1983b)). In the normal case with b unknown, the LRT statistics are $S_{01} = \sum_{i=1}^k w_i(\bar{\mu}_i-\hat{\mu})^2/\sum_{i=1}^k\sum_{j=1}^{n_i} a_i^{-2}(X_{ij}-\hat{\mu})^2$ and $S_{12} = \sum_{i=1}^k w_i(\bar{X}_i-\bar{\mu}_i)^2/\sum_{i=1}^k\sum_{j=1}^{n_i} a_i^{-2}(X_{ij}-\hat{\mu})^2$ and the tail

probabilities under H_0 are

$$pr(S_{01} \geq t) = \sum_{\ell=2}^{k} P(\ell,k;w)pr(B_{(\ell-1)/2,(N-\ell)/2} \geq t) \tag{3}$$

and

$$pr(S_{12} \geq t) = \sum_{\ell=1}^{k-1} P(\ell,k;w)pr(B_{(k-\ell)/2,(N-k)/2} \geq t), \tag{4}$$

where $N = \sum_{i=1}^{k} n_i$ and $B_{a,b}$ is a random variable with beta distribution with parameters a and b.

The level probabilities can be obtained by the explicit formulas discussed in Barlow et al. (1972, p. 146) if $k \leq 4$ and for $k > 4$ they can be obtained by their recursive relation (3.23) and repeated numerical integration of their (3.38) For the simple tree ordering and $k \geq 2$, their (3.38) and (3.23) become

$$P(k,k;w) = \int_{-\infty}^{\infty} \{\prod_{i=1}^{k-1} \Phi((\frac{w_{i+1}}{w_i})^{1/2}x)\}\phi(x)dx \tag{5}$$

and for $1 < \ell < k$,

$$P(\ell,k;w) = \sum_{B} P(\ell,\ell;W(B))P(1,k-\ell+1;w(B)), \tag{6}$$

where Φ and ϕ are the c.d.f. and p.d.f. for a standard normal distribution, the sum in (6) is over subsets of $\{1,2,\ldots,k\}$ which include 1 and have cardinality $k-\ell+1$, $W(B)$ is a vector of length ℓ whose first coordinate is $\sum_{i \in B} w_i$ and its remaining coordinates are the w_i with $i \notin B$, and $w(B)$ is a vector of length $k-\ell+1$ made up of the w_i with $i \in B$ and w_1 is its first coordinate. For the simple tree ordering, $P(\ell,k;w) = P(\ell,k;w')$ if $w_1' = w_1$ and $w_2',\ldots,w_k'$ is some permutation of $w_2,\ldots,w_k$. So the order of the last $\ell-1$ coordinates of $W(B)$ and the order of the last $k-\ell$ coordinates of $w(B)$ is not important.

Even for moderate k, computation via (5) and (6) can be quite tedious. If $w_1 = \ldots = w_k$, we suppress the w in the notation, and note that their equations (3.38) and (3.39) give

$$P(k,k) = \int_{-\infty}^{\infty} (\Phi(x))^{k-1}\phi(x)dx = 1/k; \text{ for } 2 \le \ell \le k-1$$

$$P(\ell,k) = \binom{k-1}{\ell-1} P(1,k-\ell+1)\int_{-\infty}^{\infty} [\Phi(x(k-\ell+1)^{-1/2})]^{k-1}\phi(x)dx; \qquad (7)$$

and $P(1,k) = 1 - \sum_{\ell=2}^{k} P(\ell,k)$.

The $P(\ell,k)$ for $2 \le k \le 12$ are given in Table A.6 of Barlow et al. (1972).

Since the χ^2 and beta distributions in (1) and (3) increase stochastically with ℓ, upper and lower bounds for these tail probabilities are found by obtaining upper and lower stochastic bounds for the discrete distributions, $P_k(w)$, which assign probability $P(\ell,k;w)$ to ℓ for $\ell = 1,2,\ldots,k$; cf. Theorem 1 of Robertson and Wright (1982). In (2) and (4) the χ^2 and beta distributions decrease with ℓ and so upper bounds for these tail probabilities are obtained by finding lower stochastic bounds for $P_k(w)$ and vice a versa.

Let $I_A(\cdot)$ denote the indicator function of the set A, for $k \ge 2$ and $1 \le \ell \le k$, define

$$A(\ell,k) = I_{\{k-1,k\}}(\ell)/2 \text{ and } B(\ell,k) = \binom{k-1}{\ell-1}2^{1-k}$$

and let A_k and B_k be the corresponding discrete distribution on $\{1,\ldots,k\}$. The mixing coefficients, $B(\ell,k)$ were encountered in Dykstra and Robertson (1982, 1983) and Robertson and Wright (1982). The proof of the following theorem uses an algorithm for computing the MLEs restricted by the simple tree ordering

which is an immediate consequence of either Thompson's minimum violator algorithm or the minimum lower sets algorithm (MLSA), cf. Barlow et al. (1972, Section 2.3). Arrange the treatment sample means in increasing order, $\bar{X}_{(2)} \leq \cdots \leq \bar{X}_{(k)}$, let $\bar{X}_1' = \bar{X}_1$, let $\bar{X}_i' = \bar{X}_{(i)}$ for $i = 2,\ldots,k$, let $w_1' = w_1$ and let w_i' be the weight associated with $\bar{X}_{(i)}$ for $i = 2,\ldots,k$. Choose j to be the smallest positive integer with

$$M_j = (\textstyle\sum_{i=1}^{j} w_i' \bar{X}_i')/\sum_{i=1}^{j} w_i' < \bar{X}_{j+1}'$$

and if no such j exists, let $j = k$. The restricted MLEs are $\bar{\mu}_i = M_j$ if $\bar{X}_i$ is included in M_j and $\bar{\mu}_i = \bar{X}_i$ if not. The first level set, S_k, consists of those is for which $\bar{\mu}_i = \bar{\mu}_1$ and, assuming there are no ties among $\bar{X}_{(j+1)},\ldots,\bar{X}_{(k)}$, there are $k - j + 1$ level sets.

THEOREM 1. For $k \geq 2$ and any vector, w, of positive weights, A_k and B_k are upper and lower stochastic bounds for $P_k(w)$. Furthermore, if $w(n) = (1/n,1,\ldots,1)$ and $w'(n) = (n,1,\ldots,1)$, then $P(\ell,k;w(n)) \to A(\ell,k)$ and $P(\ell,k;w'(n)) \to B(\ell,k)$ as $n \to \infty$.

Proof. To establish that A_k is an upper bound we only need to show that $P(k,k;w) \leq 1/2$. This follows immediately from $P(k,k;w) = pr(\bar{X}_1 < \bar{X}_i$ for $i = 2,\ldots,k)$. The proof that B_k is a lower bound is by induction. Clearly, $P(1,2;w) = P(2,2;w) = 1/2$. So we consider $k \geq 2$ and define N_k, S_k and $\bar{\mu}_{1,k}$ to be the number of level sets, the first level set and the value of $\bar{\mu}$ on the first level set in the restricted MLEs based on $\bar{X}_1,\ldots,\bar{X}_k$ with weights $w_1,\ldots,w_k$. Let N_{k+1}, S_{k+1} and $\mu_{1,k+1}$ be the corresponding values determined by the MLEs based on $\bar{X}_1,\ldots,\bar{X}_{k+1}$ with weights $w_1,\ldots,w_{k+1}$. As in the earlier discussion, let j_k = card. (S_k).

Define $Av(C) = \sum_{\{i:i\in C\}} w_i \bar{X}_i / \sum_{\{i:i\in C\}} w_i$ for $\phi \neq C \subset \{1,\ldots,k\}$. We first show that $N_{k+1} \geq N_k + I_{(\bar{\mu}_{1,k},\infty)}(\bar{X}_{k+1})$. Because the estimates are restricted MLEs, we know that if $\bar{X}_{k+1} > \bar{\mu}_{1,k}$, then $\bar{\mu}_{k+1,k+1} = \bar{X}_{k+1}$ and $N_{k+1} = N_k + 1$. If $\bar{X}_{k+1} \leq \bar{\mu}_{1,k}$, then we will show that $S_{k+1} \subset S_k \cup \{k+1\}$. For it not, there is an $i \in S_{k+1} - (S_k \cup \{k+1\})$ and so $i \neq 1$, $\bar{X}_{k+1} \leq Av(S_k) < \bar{X}_i$ which implies $Av(S_k \cup \{k+1\}) < \bar{X}_i$. But by the MLSA, $Av(S_{k+1}) \leq Av(S_k \cup \{k+1\})$ and hence, $Av(S_{k+1} - \{i\}) < Av(S_{k+1})$ which contradicts the choice of S_{k+1} according to the MLSA.

So $pr(N_{k+1} \geq \ell + 1) \geq pr(N_k = \ell, \bar{X}_{k+1} > \bar{\mu}_{1,k}) + pr(N_k \geq \ell + 1)$,

but the proof given for (3.23) in Barlow et al. (1972) shows that $pr(N_k = \ell, \bar{X}_{k+1} > \bar{\mu}_{1,k}) =$

$$\sum pr(Av(S_k) < \bar{X}_\alpha \text{ for all } \alpha \in \{1,\ldots,k\} - S_k, \bar{X}_{k+1} > Av(S_k))P(S_k)$$

where the summation is overall $S_k \subset \{1,\ldots,k\}$ which contain $\{1\}$ and have cardinality $k - \ell + 1$ and $P(S_k)$ is the probability that $Av(S) \geq Av(S_k)$ for all S which are proper subsets of S_k and contain $\{1\}$. But because $Av(S_k) \leq \bar{X}_1$ this sum is bounded below by

$$\tfrac{1}{2}\sum pr(Av(S_k) < \bar{X}_\alpha \text{ for all } \alpha \in \{1,\ldots,k\} - S_k)P(S_k),$$

which according to their (3.23) is $P(\ell,k;w)/2$. Applying the induction hypothesis, $pr(N_{k+1} \geq \ell + 1) \geq (pr(N_k \geq \ell) + pr(N_k \geq \ell+1))/2$ $\geq 2^{-k} \sum_{\alpha=\ell+1}^{k+1} \binom{k}{\alpha-1} = \sum_{\alpha=\ell+1}^{k+1} B(\alpha,k+1)$.

It remains to be shown that these bounds are sharp, however these results are special cases of Theorem 2 and 4. The proof is completed.

To give some idea of the influence of w on the $\bar{\chi}^2$ tail probabilities given in (1), the values of

$$\bar{\chi}^2_A(t) = \sum_{\ell=2}^{k} A(\ell,k)\,pr(\chi^2_{\ell-1} \geq t) \text{ and } \bar{\chi}^2_B(t) = \sum_{\ell=2}^{k} B(\ell,k)\,pr(\chi^2_{\ell-1} \geq t)$$

with $t = e_{0.05}(k)$, the equal weights 0.05 critical value, are given in Table 1 for various k.

3. THE EQUAL WEIGHTS APPROXIMATION

Primarily, we study the accuracy of approximations as they apply to (1), the tail probabilities of Bartholomew's $\bar{\chi}^2$ test statistic. If one uses the 0.05 equal weights critical value when in fact the weights are not equal the true significance level, α, may be quite different from 0.05. For instance, we see from Table 1, if k=5 then α may be as small as 0.0141 or as large as 0.0794. However, in practical situations the discrepancy may be much less. Since the upper and lower bounds occur as the limiting cases for $w = (1,R,\ldots,R)$ and $w = (R,1,\ldots,1)$ respectively, it is of interest to consider the corresponding discrepancies for moderate values of R. The second and fourth columns of Table 2 contain the values of (1) for these weights with $k = 5$ and $t = e_{0.05}(5) = 7.653$. We denote the value corresponding to $(1,R,\ldots,R)$ by $\bar{\chi}^2_{A(R)}(t)$ and that corresponding to $(R,1,\ldots,1)$ by $\bar{\chi}^2_{B(R)}(t)$. The results presented in Table 2 seem to indicate that if k=5 and R, the ratio of the largest weight to the smallest, is two or smaller and one uses $e_{0.05}(5)$, then $0.044 \leq \alpha \leq 0.057$ and if $R \leq 3$, then $0.040 \leq \alpha \leq 0.060$. To further substantiate these claims 5000 sets of random weights were generated. In particular, five independent weights uniform on $(1,R_u)$ were generated and the value of (1) was computed at $t = 7.653$. In 5000 replications, all the values of (1) were between 0.044 and 0.056 for $R_u = 2$, and for $R_u = 3$ they were between 0.040 and 0.060.

To give an indication of the behavior of the equal-weights approximation for arbitrary weights, random weights were generated. Five independent uniform (0,1) random numbers, $U_1,\ldots,U_5$, were generated and with $w_i = U_i$, $\bar{\chi}^2_w(7.653)$ and R were computed. The second and third columns of Table 3 contain a frequency distribution for the 5000 replications of this experiment along with the minimum value of R for each cell. For those 5000 weight sets, we note that if $R < 2.06$ then $0.044 \leq \alpha \leq 0.056$ and if $R < 3.34$ then $0.04 \leq \alpha \leq 0.06$.

Next we consider the accuracy of the approximation as k increases. While it is difficult to compute the level probabilities for arbitrary w and larger k, the $P(\ell,k;w)$, with weights of the form $B(R) = (R,1,\ldots,1)$ and $A(R) = (1,R,\ldots,R)$, are much more tractable. For $k = 8$ and $w = B(R)$, $A(R)$ the level probabilities were computed and (1) was evaluated at $t = e_{0.05}(8) = 11.807$. These results are found in the sixth and eighth columns of Table 2. While the limiting values are considerably more extreme at $k = 8$, the increase in the errors for the equal-weights approximation for $k = 8$ over $k = 5$ are not so large when $R = 2$ or 3.

For many practical situations the equal-weights approximation to the critical values of Bartholomews $\bar{\chi}^2$ test is adequate. However, in some situations, in particular for moderate k and considerable variation in the w_is, other approximations need to be considered. In the next section an approximation based on the pattern of large and small weights is considered.

4. AN APPROXIMATION BASED ON PATTERNS IN THE WEIGHTS

In the totally ordered case, Chase (1974) developed an approximation for the situation in which $w_2 = \ldots = w_k$ and $w_1/w_2 > 1$. His approximate critical values are obtained by interpolating between the critical values for the equal-weights $\bar{\chi}^2$ distribution and the $\bar{\chi}^2$ distribution corresponding to the limiting $P(\ell,k;w)$ as $w_1 \to \infty$ with $w_2 = \ldots = w_k$ fixed. Robertson and Wright (1983a) extended Chase's ideas to provide an approximation for the totally ordered case. We use a similar approach for the simple tree ordering.

First, we obtain the limiting values for $P(\ell,k;w)$ for each situation in which the w_i assume only two values, one of which remains constant and the other approaches infinity. Since $P(\ell,k;w) = P(\ell,k;w')$ for $w' = (cw_1,\ldots,cw_k)$ with $c > 0$, the exact value of the constant weights is immmaterial. The cases in which the first weight is either small or large are different.

For the case in which the first weight is large we define $Q(\ell,k;s)$ to be $\lim_{n\to\infty} P(\ell,k;w(n,s))$ with the first k-s coordinates of $w(n,s)$ equal to n and the last s coordinates equal to 1 with $1 \leq s < k$. For notational convenience we set $P(1,1) = 1$.

THEOREM 2. For $k \geq 2$ and $1 \leq s < k$, $\{Q(\ell,k;s)\}$ is the convolution of $\{P(\ell,k-s)\}$ with the binomial probabilities with parameters s and 1/2.

Proof. The proof is by induction. For $\ell = 1$ or $2, Q(\ell,2;1) = \lim_{n\to\infty} P(\ell,2;(n,1)) = .5$ and so the desired result holds for $k = 2$.

Fix $k > 2$, $2 \leq \ell \leq k$ and $1 \leq s \leq k-2$. For the vector of weights $w(n,s)$ and the simple tree ordering, (6) gives

$Q(\ell,k;s) =$

$$\lim_{n\to\infty} \sum_{\nu=0}^{s} \binom{s}{\nu}\binom{k-1-s}{k-\ell-\nu} P(\ell,\ell;w'(n,\nu))P(1,k-\ell+1;w''(n,\nu)) \qquad (8)$$

where $w'(n,\nu)$ has its first coordinate equal to $n(k-\ell-\nu+1)+\nu$, the next $\ell+\nu-1-s$ coordinates are n and the last $s-\nu$ are 1 and $w''(n,\nu)$ has its first $k-\ell-\nu+1$ coordinates equal to n and its last ν are 1. The summand is zero if $\ell+\nu-1-s$ or $k-\ell-\nu$ is negative. Applying (5),

$$\lim_{n\to\infty} P(\ell,\ell;w'(n,\nu)) = \left(\tfrac{1}{2}\right)^{s-\nu}\int_{-\infty}^{\infty} (\Phi(\lambda x))^{\ell+\nu-1-s}\phi(x)dx$$

where $\lambda = (k-\ell-\nu+1)^{-\frac{1}{2}}$, but applying (5) again the right hand side can be written as $P(\ell+\nu-s,\ell+\nu-s;(k-\ell-\nu+1,1,\ldots,1))/2^{s-\nu}$. By the inductive hypothesis, if $\nu>0$, $P(1,k-\ell+1;w''(n,\nu))$ approaches $P(1,k-\ell-\nu+1)/2^{\nu}$ and if $\nu=0$, the two expressions are equal. So (8) becomes

$$\sum_{\nu=0}^{s} \binom{s}{\nu}\left(\tfrac{1}{2}\right)^{s}\binom{k-s-1}{k-\ell-\nu} P(\ell+\nu-s,\ell+\nu-s;(k-\ell-\nu+1,1,\ldots,1))P(1,k-\ell-\nu+1)$$

and applying (3.39) of Barlow et al. (1972), this becomes

$$\sum_{\nu=0}^{s} \binom{s}{\nu}\left(\tfrac{1}{2}\right)^{s} P(\ell+\nu-s,k-s),$$

which is the desired result. Since $Q(1,k;s) = 1 - \sum_{\ell=2}^{k} Q(\ell,k;s)$ the result also holds for $\ell=1$. The proof for $s = k-1$ is like the above except (8) becomes

$$\binom{k-1}{k-\ell} P(\ell,\ell;(n-k-\ell,1,\ldots,1))P(1,k-\ell+1;(n,1,\ldots,1)).$$

The proof is completed.

For the case in which the first weight is small, we define $R(\ell,k;s)$ to be the $\lim_{n\to\infty} P(\ell,k;w^*(n,s))$ with $1 \leq s < k$, the first s coordinates of $w^*(n,s)$ equal to 1 and the other $k-s$

coordinates equal to n. We need not consider $s = k$, for if all the coordinates of $w^*(n,s)$ are 1, then $P(\ell,k;w^*(n,s)) = P(\ell,k)$. For the purpose of computing $P(\ell,k;w)$ we may assume $\mu_i = 0$ for $i = 1,\ldots,k$. Consider $w^*(n,s)$, a fixed point in the underlying probability space where the $\overline{X}_i$ are all distinct and sufficiently large n. Choose α so that $\overline{X}_\alpha = \min_{s+1\le i\le k}\overline{X}_i$. As $n \to \infty$, $\overline{X}_\alpha$ becomes degenerate at zero and so $\overline{\mu}_1 \le 0$. Since $\overline{X}_\alpha < \overline{X}_i$ for $s < i \le k$ with $i \ne \alpha$, $\overline{\mu}_1 < \overline{\mu}_i$ for such i. So the number of level sets is $k-s-1$ plus the number of distinct values in $(\overline{\mu}_1,\ldots,\overline{\mu}_s,0)$.

For $k \ge 2$ and $1 \le \ell \le k$, define $S(\ell,k) = \lim_{n\to\infty} P(\ell,k;(1,\ldots,1,n))$

<u>THEOREM 3</u>. For $k \ge 2$ and $1 \le \ell \le k$, $S(\ell,k)$ is determined by

$S(k,k)=(1-(\frac{1}{2})^{k-1})/(k-1)$, $S(1,k)=1-\sum_{\ell=2}^{k}S(\ell,k)$ and for $2\le\ell\le k-1$

$$S(\ell,k)=\binom{k-2}{k-\ell-1}S(1,k-\ell+1)/2^{\ell-1}+\binom{k-2}{k-\ell}P(1,k-\ell+1)\int_0^\infty(\Phi((k-\ell+1)^{-\frac{1}{2}}x))^{\ell-2}\phi(x)dx \tag{9}$$

<u>Proof</u>. The proof is by induction. Applying (5) and letting $n \to \infty$, we see that

$$S(k,k) = \int_0^\infty(\Phi(x))^{k-2}\phi(x)dx = (1-(\tfrac{1}{2})^{k-1})/(k-1),$$

and of course $\sum_{\ell=1}^{k} S(\ell,k) = 1$. So the desired result holds for $k=2$. Consider $k\ge3$ and $2\le\ell\le k-1$, applying (6),

$$S(\ell,k)=\lim_{n\to\infty}\{\binom{k-2}{k-\ell-1}P(\ell,\ell;(n+k-\ell,1,\ldots,1))P(1,k-\ell+1;(1,\ldots,1,n))$$

$$[illegible](\ell,\ell;(k-\ell+1,1,\ldots,1,n))P(1,k-\ell+1)\}.$$

By (5) lim [illegible];$(n-k-\ell,1,\ldots,1))=(\frac{1}{2})^{\ell-1}$ and $\lim_{n\to\infty}P(\ell,\ell;(k-\ell+1,1,\ldots,1,n)) = \int_0^\infty(\Phi((k-\ell+1)^{-1/2}x))^{\ell-2}\phi(x)dx$. The conclusion follows from the inductive hypothesis.

The values of $S(\ell,k)$ for $k=2,\ldots,10$ and $1\le\ell\le k$ are given in Table 4.

<u>THEOREM 4</u>. Let $k\ge 2$ and $1\le s<k$. For $k-s\le\ell\le k$, $R(\ell,k;s) = S(\ell+s+1-k,s+1)$ and for $1\le\ell<k-s$, $R(\ell,k;s) = 0$.

<u>Proof</u>. The proof is by induction. For $k=2$, $s=1$ and $R(\ell,2;s) = S(\ell,2)$ by definition. Consider $k\ge 3$ and $1\le s<k$ and note that because $R(\ell,k;s)$ and $S(\ell,s+1)$ are probability vectors, we only need to establish the first claim. Computing $P(k,k;w^*(n,s))$ using (5) and letting $n\to\infty$, we see that $R(k,k;s) = (1-(\frac{1}{2})^s)/s = S(s+1,s+1)$. For $\max(2,k-s)\le\ell<k$, we use (6) to compute $P(\ell,k;w^*(n,s))$ and write $R(\ell,k;s)$ as

$$\lim_{n\to\infty}\sum_{\nu=0}^{s-1}\binom{s-1}{\nu}\binom{k-s}{k-\ell-\nu}P(\ell,\ell;w^{**}(n,\nu))P(1,k-\ell+1;w^{***}(n,\nu)) \tag{10}$$

where the first coordinate of $w^{**}(n,\nu)$ is $\nu+1+(k-\ell-\nu)n$, the next $s-\nu-1$ coordinates are 1, the last $\ell-s+\nu$ coordinates are n, the first $\nu+1$ coordinates of $w^{***}(n,\nu)$ are 1, and the last $k-\ell-\nu$ coordinates are n. First consider $k-s<\ell<k$ or $0<k-\ell<s$. If $\nu>k-\ell$ then $k-\ell-\nu<0$ and the corresponding term in (10) is zero. If $\nu=k-\ell$, then $P(1,k-\ell+1;w^{***}(n,\nu)) = P(1,k-\ell+1)$ and computing $P(\ell,\ell;w^{**}(n,\nu)$ using (6) we see that the term with $\nu=k-\ell$ in (10) approaches

$$\binom{s-1}{k-\ell}P(1,k-\ell+1)\int_0^\infty(\Phi((k-\ell+1)^{-1/2}x))^{s+\ell-k-1}\phi(x)dx,$$

which is the second term in the representation of $S(\ell+s+1-k,s+1)$ given in (9). If $\nu=k-\ell-1$, then $P(1,k-\ell+1,w^{***}(n,\nu))\to S(1,k-\ell+1)$ and

$$P(\ell,\ell;w^{**}(n,\nu))\to(\tfrac{1}{2})^{\ell+s-k}\int_{-\infty}^{\infty}(\Phi(x))^{k-s-1}\phi(x)dx = (2^{\ell+s-k}(k-s))^{-1}.$$

Hence, the term with $\nu = k-\ell-1$ in (10) approaches

$$\binom{s-1}{k-\ell-1}\left(\frac{1}{2}\right)^{\ell+s-k} S(1,k-\ell+1)$$

which is the first term in the representation of $S(\ell+s+1-k,s+1)$ given in (9). For $\nu<k-\ell-1$, $P(1,k-\ell+1;w^{***}(n,\nu))\to R(1,k-\ell+1;\nu+1)$ and $k-\ell+1<s+1\le k$ and so the inductive hypothesis holds. $R(1,k-\ell+1;\nu+1) = 0$ since $1<k-\ell+1-(\nu+1) = k-\ell-\nu$.

Finally we consider $\ell=k-s$. If $s=k-1$, then $w^*(n,s) = (1,\ldots,1,$ and $R(\ell,k;k-1) = S(\ell,k) = S(1,s+1)$ which is the desired conclusion. If $s<k-1$, then (10) is applicable for $\ell=k-s\ge 2$. However, the term with $\nu=s-1$ in (10) is $(k-s)P(k-s,k-s;(s+n,n,\ldots,n))R(1,s+1)$ which approaches $(k-s)P(k-s,k-s)R(1,s+1)=R(1,s+1)$. For $\nu<s-1$, $k-\ell-\nu=s-\nu>1$ and so $P(1,k-\ell+1;w^{***}(n,\nu))\to R(1,k-\ell+1;s-\nu)$, which by the inductive hypothesis is zero. The proof is completed.

We now describe a method of obtaining approximate p-values or approximate critical values. Given a weight set $w_1,\ldots,w_k$, classify each weight, w_i, as small if $w_i\le \tilde{w}$ or large if $w_i>\tilde{w}$. We considered choices for $\tilde{w}$ of the form $\tilde{w} = \alpha \min w_i+(1-\alpha)\max w_i$ and $\tilde{w} = \beta \min w_i$ and based on numerical studies recommend $\tilde{w} = 1.5 \min w_i$. Next find the p-value or critical value determined by the limiting level probabilities if the small weights were fixed at one and the large weights approached infinity. Finally, interpolate between the limiting value and the equal-weights value. For instance, the approximate p-value is given by

$$\tilde{P} = \sqrt{r}\, P_e + (1-\sqrt{r})P_\ell \tag{11}$$

where r is average of the small weights divided by the average of the large weights, P_e is the equal-weights p-value and P_ℓ is the appropriate limiting p-value. The choice of $\frac{1}{2}$ for the exponent on r is based on numerical studies also. Replacing the p-values in (11) by the appropriate critical values, one obtains approximate critical values, which we denote by $a_\alpha(k)$.

Robertson and Wright (1983 a,b) studied the errors for analogous approximate p-values as well as the errors in the level of significance corresponding to approximate critical values. In both cases, it was found that for p-values and levels of significance of similar magnitudes, the errors were similar in size. For that reason, we only study the errors in significance levels arising from approximate critical values. Columns 3,5,7 and 9 of Table 2 contains the true significance levels corresponding to $a_{0.05}(k)$ for w = (R,1,...,1) and (1,R,...,R) with k = 5,8 and several values of R. For such weights, the approximation given in this section clearly outperforms the equal-weights approximation. It is also clear from Table 2, that the interpolation scheme given here puts too much weight on the limiting critical values for small values of R. However, this scheme was chosen because of its performance on randomly generated weight sets.

Because of the way it was developed, we would expect that this approximation would perform well for weight sets with only two distinct values and so we wish to compare it with

the equal-weights approximation for arbitrary weight sets. Recall, that Table 3 contains a frequency distribution of $\bar{\chi}_w^2(e_{0.05})$ for 5000 sets of weights which are pseudo independent, uniform random variables on (0,1). For these same weight sets the frequency distribution of $\bar{\chi}_w^2(a_{0.05})$ is also given in Table 3. It is clear that the approximation proposed in this section is also more accurate than the equal-weights approximation for these randomly generated weight sets.

REFERENCES

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). Statistical Inference Under Order Restriction. New York: Wiley.

Chase, G. R. (1974). On testing for ordered alternatives with increased sample size for a control. Biometrika 61, 569-78.

Dunnett, C. W. (1955). A multiple comparisons procedure for comparing several treatments with a control. J. Am. Statist. Assoc. 50, 1096-121.

Dunnett, C. W. (1964). New table for multiple comparisons with a control. Biometrics 20, 482-91.

Dykstra, R. L. and Robertson, T. (1982). Order restricted statistical tests on multinomial and Poisson parameters: the star shaped restriction. Ann. Statist. 10, 1246-52.

Dykstra, R. L. and Robertson, T. (1983). On testing monotone tendencies. J. Am. Statist. Assoc. 78, 342-50.

Grove, D. M. (1980). A test of independence against a class of ordered alternatives in a 2xC contingency table. J. Am. Statist. Assoc. 75, 454-9.

Magel, Rhonda and Wright, F. T. (1984). Tests for and against trends among Poisson intensities. Inequalities in Statistics and Probability: IMS Lecture Notes-Monograph Series 5 (to appear).

Robertson, T. (1978). Testing for and against an order restriction on multinomial parameters. J. Am. Statist. Assoc. 73, 197-202.

Robertson, T. and Wegman, E. J. (1978). Likelihood ratio tests for order restrictions in exponential families. Ann. Statist. 6, 485-505.

Robertson,T. and Wright, F. T. (1982). Bounds on mixtures of distributions arising in order restricted inference. Ann. Statist. 10, 302-6.

Robertson, T and Wright, F. T. (1983a). On approximation of the level probabilities and associated distributions in order restricted inference. Biometrika 70, 597-606.

Robertson, T. and Wright, F. T. (1983b). One-sided comparisons for treatments with a control. Technical Report #78, Department of Statistics; University of Iowa.

Shirley, E. (1977). A non-parametric equivalent of Williams test for contrasting increasing dose levels of a treatment. Biometrics 33, 386-9.

Siskind, V. (1976). Approximate probability integrals and critical values for Bartholomew's test for ordered means. Biometrika 63, 647-54.

Table 1. $\bar{\chi}_A^2(t)$ and $\bar{\chi}_B^2(t)$ with $t = e_{0.05}(k)$

	k							
	3	4	5	6	7	8	9	10
$\bar{\chi}_B^2(t)$	0.0331	0.0217	0.0141	0.0094	0.0058	0.0037	0.0023	0.0015
$\bar{\chi}_A^2(t)$	0.0669	0.0746	0.0794	0.0827	0.0850	0.0863	0.0881	0.0891

Table 2. $\bar{\chi}^2_{A(R)}(t)$ and $\bar{\chi}^2_{B(R)}(t)$ with $t=e_{0.05}(k)$, $a_{0.05}(k)$ for $k=5$,

	$k=5$				$k=8$			
	$w=(1,R,\ldots,R)$		$w=(R,1,\ldots,1)$		$w=(1,R,\ldots,R)$		$w=(R,1,\ldots,1)$	
R	$t=e_{0.05}(5)$	$t=a_{0.05}(5)$	$t=e_{0.05}(5)$	$t=a_{0.05}(5)$	$t=e_{0.05}(8)$	$t=a_{0.05}(8)$	$t=e_{0.05}(8)$	$t=a_{0.05}(8)$
1.5	0.0539	0.0494	0.0462	0.0506	0.0553	0.0498	0.0446	0.0528
2.0	0.0566	0.0492	0.0436	0.0506	0.0590	0.0499	0.0408	0.0536
3.0	0.0601	0.0492	0.0403	0.0501	0.0636	0.0501	0.0360	0.0534
5.0	0.0641	0.0494	0.0370	0.0491	0.0687	0.0503	0.0307	0.0520
10.0	0.0684	0.0497	0.0337	0.0481	0.0740	0.0505	0.0255	0.0495
100.0	0.0759	0.0500	0.0301	0.0482	0.0829	0.0504	0.0193	0.0472
∞	0.0794	0.0500	0.0141	0.0500	0.0863	0.0500	0.0037	0.0500

Table 3. Frequency distribution for $\bar{\chi}_w^2(7.653)$ and $\bar{\chi}_w^2(a_{0.05}(5))$ for 5000 randomly generated weight sets

	$\bar{\chi}_w^2(7.653)$		$\bar{\chi}_w^2(a_{0.05}(5))$	
Interval	Frequency	Minimum R	Frequency	Minimum R
(0.032, 0.036)	47	8.87	0	--
(0.036, 0.040)	397	4.85	3	15.41
(0.040, 0.044)	957	2.29	120	5.96
(0.044, 0.048)	1196	1.39	829	1.39
(0.048, 0.052)	868	1.14	2553	1.14
(0.052, 0.056)	543	1.39	345	2.77
(0.056, 0.060)	397	2.06	140	4.32
(0.060, 0.064)	242	3.34	54	6.88
(0.064, 0.068)	180	5.87	16	10.38
(0.068, 0.072)	110	12.11	2	112.14
(0.072, 0.076)	55	26.77	0	--
(0.076, 0.080)	8	123.94	0	--

Table 4. Limiting level probabilities $S(\ell,k)=\lim_{n\to\infty}P(\ell,k;(1,\ldots,1,n))$.

ℓ \ k	2	3	4	5	6	7	8	9	10
1	.5000	.1250	.0270	.0052	.0009	.0002	.0000	.0000	.0000
2	.5000	.5000	.2083	.0625	.0153	.0032	.0006	.0001	.0000
3		.3750	.4730	.2604	.0974	.0285	.0070	.0015	.0003
4			.2917	.4375	.2910	.1280	.0431	.0120	.0029
5				.2344	.4017	.3073	.1535	.0579	.0178
6					.1937	.3687	.3145	.1740	.0721
7						.1641	.3395	.3161	.1902
8							.1418	.3139	.3141
9								.1245	.2917
10									.1109

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A140684	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Approximating the Level Probabilities in Order Restricted Inference: The Simple Tree Ordering		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER N000 14-80-C-0322
7. AUTHOR(s) F. T. Wright and Tuan Tran		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Mathematics and Statistics, University of Missouri-Rolla, Rolla, MO 65401		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research, Department of the Navy, Arlington, VA 22217		12. REPORT DATE April 1984
		13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Order restricted inference, comparisons with a control, simple tree orderings, level probabilities, chi-bar-squared distributions

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Distribution theory for the likelihood ratio statistics for the comparison of several treatments with a control is discussed. These test statistics account for prior information that the treatments are at least as effective as the control. Their null distributions are mixtures of chi-squared or beta distributions and the mixing coefficients, which are "level probabilities", are intractable for even a moderate number of treatments and

unequal weights, which are typically the sample sizes. The distribution corresponding to equal weights is considered as an approximation and an approximation based on the pattern of large and small weights is developed. Both are adequate for moderate variations in the weights, but the approximation based on patterns in the weight set is considerably more accurate for larger variations.

8

DTIC